Administrative Codes Inaccurately Identify Recurrent Venous Thromboembolism: The CVRN VTE Study

Christine Baumgartner, MD, MAS[a,b]

Alan S. Go, MD[c,d]

Dongjie Fan, MSPH[c]

Sue Hee Sung, MPH[c]

Daniel M. Witt, Pharm D[e]

John R. Schmelzer, PhD[f]

Marc S. Williams, MD[g]

Steven H. Yale, MD[h]

Jeffrey J. VanWormer, PhD[i]

Margaret C. Fang, MD, MPH[b]

---

[a]Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse, CH-3010 Bern, Switzerland. christine.baumgartner@insel.ch

[b]University of California, San Francisco, 533 Parnassus Ave., Box 0131, room U135, San Francisco, CA 94143. Margaret.Fang@ucsf.edu

[c]Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612-2304. Alan.S.Go@kp.org; Dongjie.J.Fan@kp.org; sue.hee.sung@kp.org

[d]Departments of Epidemiology, Biostatistics and Medicine, University of California, San Francisco; Departments of Medicine, Health Research and Policy, Stanford University School of Medicine, Palo Alto, CA

[e]Department of Pharmacotherapy, University of Utah College of Pharmacy, 30 South 2000 East, Salt Lake City, UT 84112. dan.witt@pharm.utah.edu

[f]Marshfield Clinic Research Institute, 1000 North Oak Ave., Marshfield, WI 54449. Schmelzer.John@marshfieldresearch.org

[g]Genomic Medicine Institute, Geisinger, 100 N Academy Ave. Danville, PA, 17822-2620. mswilliams1@geisinger.edu

[h]Department of Medicine, University of Central Florida College of Medicine, 6850 Lake Nona Blvd., Orlando, FL 32827. Steven.Yale.md@gmail.com

[i]Marshfield Clinic Research Institute, 1000 North Oak Ave., Marshfield, WI 54449. vanwormer.jeffrey@marshfieldresearch.org

Corresponding author:

Christine Baumgartner, MD, MAS

Department of General Internal Medicine

Inselspital, Bern University Hospital

University of Bern

Freiburgstrasse

CH-3010 Bern

Switzerland

Phone: +41 (0)31 632 57 69

e-mail: christine.baumgartner@insel.ch

Funding:

This work was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (grant numbers R01HL103820, U19HL91179, and 1K24HL141354). Christine Baumgartner's work was supported by the Swiss National Science Foundation (P2BEP3_165409) and the Gottfried und Julia Bangerter-Rhyner Foundation. The sponsors did not have any role in the design and conduct of the study, collection, management, analysis, and interpretation of the data, or preparation, review or approval of the manuscript.

Conflicts of Interest: Dr. Go has received research funding through his institution from CSL Behring. Dr. Witt received research funding and consulting fees from Roche Diagnostics. None of the other authors have relationships with organizations that have a financial interest in the subject or content of the manuscript.

Dr. Fang: receives research funding from the Patient-Centered Outcomes Research Institute (PCORI award NOACs-1510-32651)

Dr. Go: research funding from the Patient-Centered Outcomes Research Institute (PCORI award NOACs-1510-32651)

All other authors have no conflicts of interest to disclose.

Text word count: 3422 (text plus tables: 4287), **Abstract word count:** 247, **References:** 30, **Tables:** 4.

Brief Title: Validity of ICD9 codes for recurrent VTE

Complete Author Information:

Christine Baumgartner, MD, MAS, Department of General Internal Medicine, Inselspital, Bern University Hospital, University of Bern, Freiburgstrasse, CH-3010 Bern, Switzerland. Phone: +41 (0)31 632 57 69; Fax: +41 (0)31 664 43 60; christine.baumgartner@insel.ch

Dongjie Fan, MSPH, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612-2304. Phone: 510-891-3565; Fax: 510-891-3508; Dongjie.J.Fan@kp.org

Alan S. Go, MD, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612-2304. Phone: 510-891-3422; Fax: 510-891-3508; Alan.S.Go@kp.org

Sue Hee Sung, MPH, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612-2304. Phone: 510-891-3476; Fax: 510-891-3508; sue.hee.sung@kp.org

Daniel M. Witt, PharmD, Department of Pharmacotherapy, University of Utah College of Pharmacy, 30 South 2000 East, Salt Lake City, UT 84112. Phone: 801-581-8851; Fax: 801-587-7923; dan.witt@pharm.utah.edu

Marc S. Williams, MD. Genomic Medicine Institute, Geisinger, 100 N Academy Ave. Danville, PA, 17822-2620. Phone 570-214-1005; Fax 570-214-7342; mswilliams1@geisinger.edu

John R. Schmelzer, PhD, MA, MS, Marshfield Clinic Research Institute, 1000 North Oak Ave., Marshfield, WI 54449. Phone: 715-389-3009; Fax: 715-389-4788; e-mail: Schmelzer.John@marshfieldresearch.org

Steven H. Yale, MD, Department of Medicine, University of Central Florida College of Medicine, 6850 Lake Nona Blvd., Orlando, FL 32827. Phone: 407-266-1000; Steven.Yale.md@gmail.com

Jeffrey J. VanWormer, PhD, Center for Clinical Epidemiology & Population Health, Marshfield Clinic Research Institute, 1000 North Oak Ave., Marshfield, WI 54449. Phone: 715-387-5241; vanwormer.jeffrey@marshfieldresearch.org

Margaret C. Fang, MD, MPH, University of California, San Francisco, 533 Parnassus Ave., Box 0131, room U135, San Francisco, CA 94143. Phone: 415-502-7100; Fax: 415-514-2094; Margaret.Fang@ucsf.edu

Administrative Codes Inaccurately Identify Recurrent Venous Thromboembolism: The CVRN VTE Study

Abstract word count: 248

Abstract

Background: Studies using administrative data commonly rely on diagnosis codes to identify venous thromboembolism (VTE) events. Our objective was to assess the validity of using International Classification of Disease, 9th Revision (ICD-9) codes in identifying recurrent VTE.

Materials and Methods: Among 5497 adults with confirmed incident VTE from four healthcare delivery systems in the Cardiovascular Research Network (CVRN), we identified all subsequent inpatient, emergency department (ED), and ambulatory clinical encounters associated with an ICD-9 code for VTE (combined with relevant radiology procedure codes for inpatient/ED VTE codes in the secondary discharge position or outpatient codes) during the follow-up period. Medical records were reviewed using standardized diagnostic criteria to assess for the presence of new, recurrent VTE. The positive predictive value (PPV) of codes was calculated as the number of valid events divided by total encounters.

Results: We identified 2397 encounters that were considered potential recurrent VTE by ICD-9 codes. However, only 31.1% (95%CI: 29.3-33.0%) of encounters were verified by reviewers as true recurrent VTE. Hospital or ED encounters with VTE codes in the primary position were more likely to represent valid recurrent VTE (PPV 61.3%, 95%CI: 56.7-66.3%) than codes in secondary positions (PPV 35.4%, 95%CI: 31.9-39.3%), or outpatient codes (PPV 20.3%, 95%CI:

18.3-22.5%). PPV was low for all VTE types (29.9% for pulmonary embolism, 38.3% for lower and 37.7% for upper extremity deep venous thrombosis, and 14.1% for other VTE).

Conclusions: ICD-9 codes do not accurately identify new VTE events in patients with a prior history of VTE.

Keywords: venous thromboembolism; pulmonary embolism; deep vein thrombosis; electronic health record; International Classification of Diseases, 9th Revision

Abbreviations: CI, confidence interval; CPT, using current procedural terminology; CVRN VTE, Cardiovascular Research Network Venous Thromboembolism; ICD-9, International Classification of Diseases, 9th Revision; KPCO, Kaiser Permanente Colorado; KPNC, Kaiser Permanente Northern California; MC, Marshfield Clinic; PPV, positive predictive value; SD, standard deviation; VTE, venous thromboembolism.

Introduction

Venous thromboembolism (VTE) (i.e., deep venous thrombosis or pulmonary embolism) has an estimated annual incidence of 900,000 events in the U.S.[1, 2] VTE poses a substantial clinical and economic burden to individuals and society, not only because of the high morbidity, mortality, and costs associated with the initial event, but also because of the risk of recurrent VTE events.[1-3] Recognition of the urgent need to optimize treatment and outcomes of VTE by the Center for Medicaid and Medicare Services, the Agency for Healthcare Research and Quality, and the Surgeon General,[2, 4, 5] has led to efforts to widespread collection of VTE-related data for quality improvement and outcomes research.

Administrative databases are increasingly used for potential quality surveillance as well as pharmacoepidemiologic and outcomes research due to multiple advantages, including their widespread availability, large population coverage, reflection of "real world" practice, lower research-related costs, and longer observation periods than prospective cohort studies of selected participants.[6] However, administrative databases have relied on coded diagnoses such as *International Classification of Disease, 9th Revision* (ICD-9) diagnosis codes with questionable validity for many health conditions.[6, 7] Even with the introduction of ICD-10 codes, which have enhanced granularity over older codes,[8] ICD-9 codes are still frequently used to identify disease conditions prior to the widespread use of ICD-10 codes and important in retrospective studies.[9-11] Prior research investigating the accuracy of VTE diagnosis codes have primarily focused on incident VTE events or have not separated out incident from prevalent events. This research has yielded widely varying results depending on the included population, with positive predictive values (PPV) as low as 26% to 31% in outpatients.[7, 12, 13]

Very little attention has been given to assessment of the validity of administrative codes to identify new VTE events in patients with a prior history of VTE (e.g., recurrent VTE events), even though recurrent VTE is one of the most important outcome measures in VTE research.

Understanding the accuracy of administrative codes to identify recurrent VTE is important to determine their utility for use in research, quality improvement efforts for secondary VTE prophylaxis, and policy-making. Therefore, we assessed the validity of ICD-9 codes for identifying recurrent VTE events against standardized criteria using medical records in a diverse, multi-institutional, community-based population of adults with previous VTE.

Material and Methods

Setting and Population

The Cardiovascular Research Network Venous Thromboembolism (CVRN VTE) Study included adults with VTE from four integrated healthcare delivery systems in the U.S. (Kaiser Permanente Northern California [KPNC], Kaiser Permanente Colorado [KPCO], Marshfield Clinic [MC], and Geisinger Health System), reflecting a diverse population in the corresponding geographic areas of Northern California, the Denver metropolitan area in Colorado, central and northwest Wisconsin, and central and northeast Pennsylvania, respectively. This study was considered exempt according to the updated Common Rule for Human Research.[14]

Details of the CVRN VTE study have been previously published.[13, 15] Briefly, health plan members aged ≥21 years with a confirmed incident VTE event from October 1, 2004 to December 31, 2010 and who had at least 12 months of continuous health plan membership with pharmacy benefits before the index VTE event were included in the study. To restrict the sample to patients with a first VTE event, the index VTE diagnosis was defined as the first

medical encounter with a diagnosis code for VTE. Patients with a VTE diagnosis code or a prescription of anticoagulants within four years before the index event were excluded. VTE diagnoses were identified from electronic medical records and based on ICD-9 codes for pulmonary embolism (415.1x), deep venous thrombosis of the lower extremities (451.1x, 451.2, 451.81, 453.4x, 453.5x), upper extremities (451.83, 451.84, 451.89, 453.72, 453.73, 453.74, 453.75, 453.76, 453.77, 453.82, 453.83, 453.84, 453.85, 453.86, 453.87) or other venous thrombosis (451, 451.9, 452, 453, 453.0, 453.1, 453.2, 453.3, 453.79, 453.8, 453.89, 453.9). ICD-9 codes were used during the entire period of the study. We did not include diagnosis codes for superficial thrombophlebitis or pregnancy-related VTE. Any inpatient or emergency department encounter with a relevant ICD-9 diagnosis code in the primary or secondary position as well as outpatient encounters were included. All administratively coded VTE events underwent chart review by trained physician and pharmacist chart abstractors using standardized criteria for validation and collection of additional information that were not otherwise available from administrative data. The final CVRN VTE study population consisted of 5,497 patients with a valid first time incident VTE event.

Identification and Validation of Recurrent VTE Events

We considered a valid recurrent VTE diagnosis to be a confirmed new pulmonary embolism or deep venous thrombosis (lower or upper extremity) occurring at the same or new site. To identify possible recurrent VTE events, we searched electronic medical records for all healthcare encounters with an ICD-9 diagnosis code for pulmonary embolism, deep venous thrombosis of the lower and upper extremities, or other venous thrombosis (e.g. hepatic vein, inferior vena cava, see **Appendix Table A.1**) subsequent to the initial VTE discharge. Follow-up

for recurrent VTE events occurred until death, disenrollment from the health system or end of administrative follow-up on December 31, 2013. Individual patients could have more than one recurrent VTE event. We included inpatient, emergency department, and outpatient encounters using the same set of ICD-9 VTE codes used for identifying incident VTE. For inpatient and emergency department encounters, relevant codes in the primary or secondary position were considered. Inpatient/emergency encounters with codes in the secondary discharge position and outpatient encounters were considered if there was a relevant radiology code present within 7 days of the diagnosis (i.e. up to 7 days prior and up to 7 days after). Relevant radiology procedure codes were identified using current procedural terminology (CPT) codes for ultrasound examinations of the extremity veins, inferior vena cava, or pelvic vessels, CPT codes for ventilation-perfusion scan of the lung, or CPT codes for computed tomography of the chest with contrast (**Appendix Table A.2**).

Medical records of all possible recurrent VTE events were manually reviewed at 3 of the 4 medical sites (KPCO, MC, Geisinger) by trained chart abstractors for validation. For the fourth site (KPNC), we used a combination of manual chart review and automated text processing (i.e., natural language processing) to verify whether events were valid recurrent VTE due to the large number of possible events. This automated text process scanned the unstructured electronic note text for mentions of recurrent VTE preceded or followed by a negation term using a clinical negation ontology to identify encounters that were considered "definitely not VTE," and these events were categorized as "invalid" VTE events. Trained physician and pharmacist reviewers using structured data forms manually reviewed encounters assessed by the automated text processing method as "possible VTE" or "likely to be VTE."

The chart review process was similar to that used in our previously described study of incident VTE events.[13] All available hospital admission, transfer and discharge records, emergency department and outpatient notes, and radiology reports within 72 hours before and after the recurrent VTE encounter were obtained and reviewed by trained reviewers. A standardized data abstraction form was used to identify valid recurrent VTE events. We determined a recurrent VTE event to be valid if there was evidence of an acute VTE event in radiology, operative, or autopsy records or as described by physician notes. If the presentation and management described in the current encounter represented a continuation of a prior episode of VTE, we did not consider the event to be valid unless there was new symptomatology, extension or progression of prior thrombus, or VTE in a new location.[16, 17] For encounters that were not considered valid recurrent VTE events, reviewers were required to indicate the reason why they were considered invalid.

Patient Characteristics

Characteristics of eligible patients were obtained from clinical and administrative databases using extracted data standardized to a Virtual Data Warehouse.[18] We also searched health plan pharmacy databases for filled anticoagulant medication prescriptions (warfarin, low molecular weight heparin, and fondaparinux) within 7 days of the possible recurrent VTE event. We did not include direct oral anticoagulants, because they had not been approved in the U.S. for treatment of acute VTE during the study period.

Statistical Analysis

All analyses were performed using SAS software version 9.1 (SAS Institute, Cary, NC).

The primary outcome was the PPV of ICD-9 codes suggesting recurrent VTE events compared to manual review of medical records. PPV was calculated as the number of chart-validated VTE events divided by the total number of events identified by ICD-9 codes with associated 95% confidence intervals. We investigated whether the PPV for recurrent VTE differed by type of VTE (i.e. pulmonary embolism, lower extremity deep venous thrombosis, upper extremity deep venous thrombosis, and other VTE), and by each individual ICD-9 code identifying any VTE event (see above). To investigate whether we could find a specific set of characteristics that can improve PPV, we also assessed whether the PPV for recurrent VTE varied depending on the discharge position of the diagnosis code (primary compared to secondary position) or encounter setting (inpatient or emergency department encounters compared to outpatient encounters). We also assessed whether recurrent VTE events could be more accurately identified if we used a combination of ICD-9 diagnosis codes with a filled prescription for anticoagulants within 7 days of the possible recurrent VTE encounter in patients that had not filled an anticoagulant prescription for at least 6 months prior to the date of the possible recurrent VTE event.

Because the first recurrent VTE event is most relevant for pharmacoepidemiologic studies, we conducted a sensitivity analysis restricting to the first encounter that met criteria during follow-up to assess whether this would improve the accuracy of identifying recurrent VTE. In a second sensitivity analysis, we implemented a lag period by excluding potential recurrent VTE events that occurred within 7 days of the incident event, as they could represent false positive findings referring to the index VTE rather than a new event.

Results

The CVRN VTE cohort included 5,497 adults who had confirmed acute incident VTE, with mean (SD) age of the cohort 65.7 (15.8) years, 51.4% women and 79.6% of white race. During the follow-up period, 2,397 eligible clinical encounters from 1,592 unique patients were assigned an ICD-9 code for VTE that represented potential acute recurrent VTE events (only considering inpatient/emergency department codes in the secondary position or outpatient codes if a relevant radiology procedure code was present within 7 days). Sixty-nine percent of patients had just one clinical encounter with a VTE code in the follow-up period, while 19.7% had 2, 6.3% had 3, and the remaining 5.1% had 4 or more. There were slightly more qualifying ambulatory encounters than inpatient/emergency department visits (**Table 1**) and most encounter codes were for pulmonary embolism (41.1%) or isolated lower extremity deep venous thrombosis (40.7%).

Predictive Value of ICD-9 Diagnosis Codes to Identify Recurrent Venous Thromboembolism

Among the 2,397 encounters with a diagnosis of VTE, 1,390 were first assessed using automated text processing, and 615 were identified as clearly invalid encounters and did not undergo further chart review validation. The remaining 1,782 medical records were manually reviewed. Overall, 746 encounters were identified as valid acute VTE events, for an overall PPV of 31.1% (95% CI: 29.3-33.0%) (**Table 2**). The PPV differed by type of VTE event, encounter setting, and whether the diagnosis code was in the primary or secondary discharge position. VTE codes from the ambulatory care setting represented acute new VTE in only 20.3% (95% CI: 18.3-22.5%) of encounters. Secondary diagnosis codes from the hospital/emergency department setting were marginally better at 35.4% (95% CI: 31.9-39.3%). For patients in the

hospital/emergency department setting, PPV for pulmonary embolism codes in the primary position was 67.0% (95% CI: 61.0-73.6%). PPVs for VTE codes of other sites were significantly lower (**Table 2**). The PPV associated with individual VTE codes are presented in the **Supplemental Table**.

In all, 596 VTE encounters were from patients who had not been prescribed anticoagulants in the prior 6 months and then had a new prescription for anticoagulant within 7 days of the clinical encounter. Incorporating anticoagulant prescription increased PPV to 54.2% (**Table 3**). VTE codes for pulmonary embolism, or of the upper or lower extremities, were reasonably accurate in the hospital/emergency department setting when coupled with anticoagulant prescription data if the diagnosis codes were in the primary position, or when considering codes in both primary and secondary positions combined (PPV 67.1% for any VTE, 73.2% for pulmonary embolism, and 67.6% for lower extremity deep venous thrombosis; results not shown). Ambulatory codes continued to be poorly predictive of actual recurrent VTE events even when incorporating new anticoagulant dispensing data (**Table 3**).

In a sensitivity analysis restricting to the 1592 first encounters with a potential recurrent VTE event, the results remained similar: the overall PPV to identify any recurrent VTE event only slightly increased to 33.7% (**Appendix Table A.3**). The exclusion of potential recurrent VTE events that occurred within 7 days of the incident event using a second sensitivity analysis did not substantially change the results (**Appendix Table A.4**).

For the 1,036 clinical encounters that underwent manual chart review and that did not represent valid recurrent VTE, reviewers attempted to ascertain the reasons why the events were not considered acute VTE events. In nearly two-thirds of all cases, the reason was that the

encounter referred to a past history of VTE but there was no evidence of new, acute event (**Table 4**).

Discussion

Within a large, diverse, community-based sample of adults with a prior VTE receiving care within four integrated healthcare delivery systems, we found that ICD-9 diagnosis codes were not accurate in identifying recurrent VTE events. The validity of the relevant ICD-9 codes varied widely based on the type of VTE, diagnostic position of the relevant code, and encounter setting. In particular, diagnosis codes representing splanchnic thrombosis or thrombotic events at unspecified locations were very rarely related to true recurrent VTE events. Because of the low overall validity of administrative codes in identifying recurrent VTE, caution should be taken when interpreting findings of studies that rely solely on administrative codes to identify recurrent VTE events. Combination of ICD-9 codes with a new prescription for anticoagulants may result in improved accuracy to identify recurrent VTE diagnosed in the hospital or emergency department setting.

Our study found that using administrative diagnosis codes for recurrent VTE events is even less accurate than when patients present with an first acute VTE episode.[13] This finding is in large part due to carry-over of past VTE codes in encounters (e.g. via patient history); although a substantial proportion of codes also represented initial suspicion for acute VTE that was subsequently ruled out after clinical evaluation. Our results are notable because while several studies have investigated the validity of ICD-9 codes for incident VTE,[7, 13, 19, 20] evidence on the validity of codes to identify *recurrent* VTE events is very limited. A previous

study that validated ICD-9 codes to identify recurrent VTE events in claims data reported a PPV of 89%.[21] That study incorporated receipt of anticoagulation into their case identification algorithm and only assessed inpatient diagnoses of recurrent VTE events, along with including only 36 potential cases of recurrent VTE leading to limited precision and potential sampling bias.[21] In contrast, we performed manual review of medical records in more than 1,000 encounters for potential recurrent VTE from both the inpatient and outpatient settings. It is also important to note that the issues that result in poor performance, in particular, carry-over from prior encounters, would not be addressed by improvements associated with the use of updated coding systems (e.g., ICD-10).

Our results imply that findings and conclusions from research and policies that used administrative codes to define recurrent VTE events without confirmatory chart review are likely to be biased. Outcome misclassification tends to bias the results towards the null in cases of non-differential misclassification, while the direction of the bias is unknown when differential misclassification is present. Our findings suggest that previous large epidemiological studies that relied only on ICD-9 codes to define recurrent VTE events[22-25] would presumably overestimate true rates of recurrent VTE events due to associated high false-positive rates, in particular for studies that include ambulatory encounters.

Prevention of VTE events is among the top priorities to improve patient safety in hospitals,[2, 5] and further research to create evidence with the aim of optimizing prevention and treatment strategies is urgently needed. While administrative databases include enormous repositories of health care-related data and facilitates comparatively efficient and more cost-effective generation of new evidence by conducting observational studies on a large scale,[26]

our results underline the need of ongoing efforts to ensure the accurate coding of diagnoses by physicians and coders, the importance of an adequate awareness of the performance of algorithms to identify health outcomes in order to correctly interpret and use research results that are based on administrative codes, and the necessity for further efforts to accurately define VTE outcomes in administrative databases. ICD-10 diagnosis codes have replaced relevant ICD-9 codes in October 2015 in the U.S, which have created more granularity in diagnosis, but a recent study shows that ICD-10 codes have similar poor accuracy for identifying thromboembolic complications of anticoagulation therapy including recurrent VTE events.[27] Our results showed that restricting inpatient diagnoses or ICD-9 codes in the primary position only does not sufficiently improve the PPV to a range that would be acceptable for outcomes research or quality improvement efforts. Combining new anticoagulant dispensing data from pharmacy databases with diagnostic codes of recurrent VTE did improve the overall PPV, resulting in a reasonable accuracy for identifying VTE recurrences if using ICD-9 codes in the primary position for hospital and emergency department encounters. However, given that only about a fourth (286 of 1036) of all hospital and emergency department encounters with potential recurrent VTE met the criteria for new anticoagulation prescriptions, there is a risk of missing true events. ICD-9 codes for outpatient encounters failed to consistently distinguish recurrent VTE events from follow-up visits of patients anticoagulated for a previous VTE despite this improvement, reflecting the lower pre-test probability of VTE in the outpatient setting. Strategies to improve the validity of outcome diagnoses in health databases could include limiting codes to those associated with the “present on admission” flag, combination of administrative data and validated automated text processing methods with incorporation of

additional laboratory and medication criteria or procedure codes, or the conduct of manual chart review to more accurately identify the diagnosis of interest.[28, 29] Quantitative bias analysis could represent another possibility to address potential misclassification resulting from imperfect validity of diagnoses.[30]

The strength of our study was that we were able to assess the validity of recurrent VTE events by performing detailed manual review of medical records on a large number of encounters within administrative databases from a spectrum of healthcare delivery systems in the U.S. We used standardized criteria and considered all available medical data for manual review including admission, transfer and discharge records, emergency department and outpatient notes, and radiology reports, implying a high accuracy of our reported results for diagnosed recurrent VTE events.

Our study has some limitations. As we did not manually screen charts without specific ICD-9 diagnosis codes suggestive of possible recurrent VTE, we were unable to determine sensitivity or specificity of ICD-9 codes to identify recurrent VTE events. We therefore are unable to draw firm conclusions on whether the use of administrative codes leads to an overestimation or underestimation of recurrent VTE events, particularly in specific subpopulations such as surgical patients.[7] However, given the low PPV, recurrent VTE events are likely overestimated. Our results cannot be directly applied to research using the new set of ICD-10 codes that were implemented starting in October 2015 nationally.[8] Our results may not be fully generalizable to all other health systems and datasets, although the four integrated healthcare delivery systems represent numerous hospitals and clinics serving a diverse population in the corresponding geographic areas. Finally, we were not able to identify the

reasons for why codes suggesting recurrent VTE events were invalid for the 615 encounters that were only assessed using automated text processing, which may differ from the reasons identified by chart reviewers.

Conclusions

In conclusion, the validity of administrative ICD-9 codes to identify recurrent VTE events is insufficient, which is particularly relevant for pharmacoepidemiologic studies or outcomes research that solely rely on administrative codes to define recurrent VTE outcomes. Our findings suggest that the results of such studies should be interpreted with caution and that strategies to account for misclassification and further efforts to improve algorithms for ascertaining recurrent VTE events from medical records for research and quality improvement purposes are of utmost importance.

Acknowledgements: None

References

[1] C.E. Mahan, Regulatory, policy and quality update for venous thromboembolism and stroke in United States hospitals, Thromb Res 130(4) (2012) 586-90.

[2] T.W. Wakefield, R.B. McLafferty, J.M. Lohr, J.A. Caprini, D.L. Gillespie, M.A. Passman, Call to action to prevent venous thromboembolism, J Vasc Surg 49(6) (2009) 1620-3.

[3] E.A. Nutescu, C. Crivera, J.R. Schein, B.K. Bookhart, Incidence of hospital readmission in patients diagnosed with DVT and PE: clinical burden of recurrent events, Int J Clin Pract 69(3) (2015) 321-7.

[4] The Centers for Medicare & Medicaid Services (CMS) and The Joint Commission, National Hospital Inpatient Quality Reporting Measures Specifications Manual, Release Notes Version: 5.0. Available at: https://www.jointcommission.org/assets/1/6/IQRManualReleaseNotes_v5_01.PDF. Accessed July 18, 2019 (2015).

[5] G. Maynard, Preventing hospital-associated venous thromboembolism: a guide for effective quality improvement, 2nd ed. Rockville, MD, Agency for Healthcare Research and Quality (2016. AHRQ Publication No. 16-0001-EF. Available at: https://www.ahrq.gov/sites/default/files/publications/files/vteguide.pdf, Accessed May 14, 2018).

[6] C. Mazzali, P. Duca, Use of administrative data in healthcare research, Internal & Emergency Medicine 10(4) (2015) 517-24.

[7] L. Tamariz, T. Harkins, V. Nair, A systematic review of validated methods for identifying venous thromboembolism using administrative and claims data, Pharmacoepidemiol Drug Saf 21 Suppl 1 (2012) 154-62.

[8] Centers for Medicare and Medicaid Services, ICD-10. Available at: https://www.cms.gov/Medicare/Coding/ICD10/index.html, (Accessed July 18, 2019).

[9] M. Mayhew, L.L. DeBar, R.A. Deyo, R.D. Kerns, J.L. Goulet, C.A. Brandt, M. Von Korff, Development and Assessment of a Crosswalk Between ICD-9-CM and ICD-10-CM to Identify Patients with Common Pain Conditions, J Pain (2019).

[10] J.A. Columbo, R. Kang, S.W. Trooboff, K.S. Jahn, C.J. Martinez, K.O. Moore, A.M. Austin, N.E. Morden, C.G. Brooks, J.S. Skinner, P.P. Goodney, Validating Publicly Available Crosswalks for Translating ICD-9 to ICD-10 Diagnosis Codes for Cardiovascular Outcomes Research, Circulation: Cardiovascular Quality & Outcomes 11(10) (2018) e004782.

[11] Patient-Centered Outcomes Research Institute, Comparing the Benefits and Harms of Medicines for Long-Term Treatment of Blood Clots -- The ALTERNATIVE Study. Available at: https://www.pcori.org/research-results/2016/comparing-benefits-and-harms-medicines-long-term-treatment-blood-clots, (Accessed July 18, 2019).

[12] F.A. Spencer, D. Lessard, C. Emery, G. Reed, R.J. Goldberg, Venous thromboembolism in the outpatient setting, Arch Intern Med 167(14) (2007) 1471-5.

[13] M.C. Fang, D. Fan, S.H. Sung, D.M. Witt, J.R. Schmelzer, S.R. Steinhubl, S.H. Yale, A.S. Go, Validity of Using Inpatient and Outpatient Administrative Codes to Identify Acute Venous Thromboembolism: The CVRN VTE Study, Med Care 55(12) (2017) e137-e143.

[14] U.S. Department of Health & Human Services, Office for Human Reserach Protections. Revised Common Rule. Available at: https://www.hhs.gov/ohrp/education-and-outreach/revised-common-rule/revised-common-rule-q-and-a/index.html., (Accessed July 19, 2019).

[15] M.C. Fang, D. Fan, S.H. Sung, D.M. Witt, J.R. Schmelzer, M.S. Williams, S.H. Yale, C. Baumgartner, A.S. Go, Treatment and Outcomes of Acute Pulmonary Embolism and Deep Venous Thrombosis: The Cardiovascular Research Network Venous Thromboembolism (CVRN VTE) Study, Am J Med (2019).

[16] J. Perez Botero, W.D. Ormsby, A.A. Ashrani, R.D. McBane, 2nd, W.E. Wysokinski, M.M. Patnaik, B.R. Lewis, D.E. Grill, R.K. Pruthi, J.A. Heit, Do incident and recurrent venous thromboembolism risks truly differ between heterozygous and homozygous Factor V Leiden carriers? A retrospective cohort study, Eur J Intern Med 30 (2016) 77-81.

[17] F.A. Spencer, C. Emery, D. Lessard, F. Anderson, S. Emani, J. Aragam, R.C. Becker, R.J. Goldberg, The Worcester Venous Thromboembolism study: a population-based study of the clinical epidemiology of venous thromboembolism, J Gen Intern Med 21(7) (2006) 722-7.

[18] A.S. Go, D.J. Magid, B. Wells, S.H. Sung, A.E. Cassidy-Bushrow, R.T. Greenlee, R.D. Langer, T.A. Lieu, K.L. Margolis, F.A. Masoudi, C.J. McNeal, G.H. Murata, K.M. Newton, R. Novotny, K. Reynolds, D.W. Roblin, D.H. Smith, S. Vupputuri, R.E. White, J. Olson, J.S. Rumsfeld, J.H. Gurwitz, The Cardiovascular Research Network: a new paradigm for cardiovascular quality and outcomes research, Circulation: Cardiovascular Quality & Outcomes 1(2) (2008) 138-47.

[19] R.H. White, M. Garcia, B. Sadeghi, D.J. Tancredi, P. Zrelak, J. Cuny, P. Sama, H. Gammon, S. Schmaltz, P.S. Romano, Evaluation of the predictive value of ICD-9-CM coded administrative data for venous thromboembolism in the United States, Thromb Res 126(1) (2010) 61-7.

[20] F. Al-Ani, S. Shariff, L. Siqueira, A. Seyam, A. Lazo-Langner, Identifying venous thromboembolism and major bleeding in emergency room discharges using administrative data, Thromb Res 136(6) (2015) 1195-8.

[21] V.J. Willey, M.F. Bullano, O. Hauch, M. Reynolds, G. Wygant, L. Hoffman, G. Mayzell, A.C. Spyropoulos, Management patterns and outcomes of patients with venous thromboembolism in the usual community practice setting, Clin Ther 26(7) (2004) 1149-59.

[22] T.H. Toledano, D. Kondal, S.R. Kahn, V. Tagalakis, The occurrence of venous thromboembolism in cancer patients following major surgery, Thromb Res 131(1) (2013) e1-5.

[23] M.B. Streiff, D. Milentijevic, K. McCrae, D. Yannicelli, J. Fortier, W.W. Nelson, F. Laliberte, C. Crivera, P. Lefebvre, J. Schein, A.A. Khorana, Effectiveness and safety of anticoagulants for the treatment of venous thromboembolism in patients with cancer, Am J Hematol 93(5) (2018) 664-671.

[24] C.I. Coleman, A.G.G. Turpie, T.J. Bunz, J. Beyer-Westendorf, Effectiveness and Safety of Rivaroxaban Versus Warfarin in Frail Patients with Venous Thromboembolism, Am J Med (2018).

[25] C.I. Coleman, T.J. Bunz, A.G.G. Turpie, Effectiveness and safety of rivaroxaban versus warfarin for treatment and prevention of recurrence of venous thromboembolism, Thrombosis & Haemostasis 117(10) (2017) 1841-1847.

[26] N. Gavrielov-Yusim, M. Friger, Use of administrative medical databases in population-based research, Journal of Epidemiology & Community Health 68(3) (2014) 283-7.

[27] K. Lawrence, C. Joos, A.E. Jones, S.A. Johnson, D.M. Witt, Assessing the accuracy of ICD-10 codes for identifying acute thromboembolic events among patients receiving anticoagulation therapy, Journal of Thrombosis & Thrombolysis 48(2) (2019) 181-186.

[28] G.S. Alotaibi, C. Wu, A. Senthilselvan, M.S. McMurtry, The validity of ICD codes coupled with imaging procedure codes for identifying acute venous thromboembolism using administrative data, Vasc Med 20(4) (2015) 364-8.

[29] K.M. Sanfilippo, T.F. Wang, B.F. Gage, W. Liu, K.R. Carson, Improving accuracy of International Classification of Diseases codes for venous thromboembolism in administrative data, Thromb Res 135(4) (2015) 616-20.

[30] T.L. Lash, M.P. Fox, R.F. MacLehose, G. Maldonado, L.C. McCandless, S. Greenland, Good practices for quantitative bias analysis, Int J Epidemiol 43(6) (2014) 1969-85.